ARI Research Note 90-32

DTIC FILE COPY

AD-A225 412

Factual Knowledge of Oregon College Students

Sarah Lichtenstein

Perceptronics, Inc.

for

Contracting Officer's Representative
Michael Drillings

Basic Research
Michael Kaplan, Director

June 1990

FIFTY YEARS OF SERVICE
1940 - 1990

United States Army
Research Institute for the Behavioral and Social Sciences

90

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

JON W. BLADES
COL, IN
Commanding

Research accomplished under contract for the Department of the Army

Perceptronics, Inc.

Technical review by

Nehama Babin

By

A-1

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS
--

2a. SECURITY CLASSIFICATION AUTHORITY
--

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE
--

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution is unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
PDRIR-1129-8-87

5. MONITORING ORGANIZATION REPORT NUMBER(S)
ARI Research Note 90-32

6a. NAME OF PERFORMING ORGANIZATION
Perceptronics, Inc.

6b. OFFICE SYMBOL *(If applicable)*
--

7a. NAME OF MONITORING ORGANIZATION
U.S. Army Research Institute for the Behavioral and Social Sciences

6c. ADDRESS *(City, State, and ZIP Code)*
6271 Variel Avenue
Woodland Hills, CA 91367

7b. ADDRESS *(City, State, and ZIP Code)*
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

8a. NAME OF FUNDING / SPONSORING ORGANIZATION U.S. Army Research Institute for the Behavioral and Social Sciences

8b. OFFICE SYMBOL *(If applicable)*
--

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
MDA903-83-C-0918

8c. ADDRESS *(City, State, and ZIP Code)*
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102B	74F		

11. TITLE *(Include Security Classification)*
Factual Knowledge of Oregon College Students

12. PERSONAL AUTHOR(S)
Lichtenstein, Sarah (Perceptronics, Inc.)

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(Year, Month, Day)*	15. PAGE COUNT
Interim	FROM 83/04 TO 87/08	1990, June	18

16. SUPPLEMENTARY NOTATION
Contracting Officer's Representative, Michael Drillings

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Algorithmic decomposition; Factual knowledge
Knowledge acquisition,
Subjective estimates

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This paper explores usefulness of algorithmic decomposition in estimating uncertain quantities. In a series of experiments, experimental questions were used that people are unlikely to know the precise answer to, but could estimate from other facts or estimates. The three studies conducted indicate that the subjects who volunteer for our experiments frequently lack knowledge of basic facts. A mild manipulation intended to persuade the subjects to reevaluate and improve their estimates had little effect. The data further suggest that efforts to develop estimation aids based on algorithmic decomposition techniques may fail, not because people are unable to make such decompositions but because they have and hold extensive stores of misinformation. When this misinformation is brought to bear on estimation problems, performance may not improve but become worse than previously expected.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE *(Include Area Code)*	22c. OFFICE SYMBOL
Michael Drillings	(202) 274-8722	PERI-br

DD Form 1473, JUN 86 *Previous editions are obsolete.*

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

FACTUAL KNOWLEDGE OF OREGON COLLEGE STUDENTS

In a series of experiments, we have been exploring the usefulness of algorithmic decomposition as an aid to estimating uncertain quantities. For our stimuli, we have been using questions that people are unlikely to know the precise answer to, but could estimate from other facts or estimates. For example, on first blush you may have no idea of how many pounds of potato chips are consumed in the U.S. in a year. But if you know that there are about 230 million people in the U.S., then an estimate of the yearly per capita consumption of chips will yield you an estimate of the requested number. The use of such a simple algorithm should lead to better performance, since an error by a factor of, say, 20 seems not unlikely for the initial question (the answer to which, at least in 1982, was 972,300,000 lbs.) but to err by a factor of 20 in estimating yearly per capita consumption (e.g., estimates of 3.4 oz. per year or 85 lbs. per year) seems less likely.

This reasoning assumes that there are some facts, such as the approximate populations of the U.S., that are generally known and known, specifically, to the University of Oregon college students who constitute the vast majority of our subject population. This paper explores that assumption.

Study 1

In a larger study, we asked subjects to estimate the populations of Oregon and of the U.S., facts that would be helpful to know in answering some of our other questions. That group of subjects showed us a considerable range of answers to

these questions, as shown in the first column of Tables 1 and 2. The subjects in that study were 245 people who answered an ad in the University of Oregon student newspaper. The ad was placed, and the experiment was run, during the University's summer session. We were concerned that summer students are atypical; in particular, the increase in the number of foreign students at the University during the summer months might bias the results. We therefore repeated the study the following fall quarter.

Insert Tables 1 & 2 about here

Study 2

The subjects were 170 people who answered an ad placed in the University of Oregon student newspaper during the University's fall quarter. The experiment was run in large campus classrooms. The subjects also completed other, unrelated tasks and were paid for their participation. For this experiment they were asked to estimate, as carefully as possible, the population of the U.S., the population of Oregon, and the number of pounds in a ton.

The results of the population questions are shown in the second column of Tables 1 and 2. The performance of these subjects was even worse than was found in Study 1. Less than 40% of the subjects gave us answers in the right "ballpark" (which we arbitrarily defined as equal to or greater than 200 million but less than 300 million for the U.S. and equal to or greater than 2

Table 1

Estimates of the Population of the U.S.

	Frequency of Responses		
Log_{10} of Estimate	Study 1	Study 2	Study 3
≥ 11.0	2	4	7
10.0 to 10.9	6	9	3
9.0 to 9.9	44	43	30
8.0 to 8.9	161	82	64
7.0 to 7.9	24	18	12
6.0 to 6.9	7	6	7
5.0 to 5.9	0	1	0
No answer	0	7	5
Number of Subjects	245	170	128
% in "ballpark"[a]	47%	32%	41%
Maximum	748 billion	300 billion	6 trillion
Third Quartile	600 million	2 billion	2 billion
Median	249 million	250 million	250 million
First Quartile	210 million	215 million	200 million
Minimum	82 thousand	220 thousand	1.5 million

[a] 200 million ≤ estimate < 300 million

Table 2

Estimates of the Population of Oregon

	Frequency of Responses		
Log_{10} of Estimate	Study 1	Study 2	Study 3
≥ 9.0	0	2	0
8.0 to 8.9	2	1	2
7.0 to 7.9	13	9	5
6.0 to 6.9	191	116	88
5.0 to 5.9	31	28	27
4.0 to 4.9	7	5	3
No answer	1	8	3
Number of Subjects	245	170	128
% in "ballpark"[a]	47%	35%	41%
Maximum	800 million	2.5 billion	250 million
Third Quartile	3 million	3.2 million	3 million
Median	2.2 million	2 million	2 million
First Quartile	1.5 million	1 million	1 million
Minimum	17 thousand	8 thousand	75 thousand

[a] 2 million ≤ estimate < 4 million

million but less than 4 million for Oregon).

One would suppose that a person who made a high estimate on the population of the U.S. would also make a high estimate of the population of Oregon. We formed the ratio of the two estimates for each subject. The results are shown in Table 3. The correct ratio is about 88. Most of our subjects showed ratios far from this. We regard as not entirely unreasonable any ratio in the inclusive range from 50 to 200: 50 would be the ratio if Oregon were average in population across the states whereas 200 would suggest a very unpopulated state relative to the others. Only 37% of our subjects showed ratios in this range. Two subjects gave populations for Oregon larger than their U.S. estimates. One can assume that these were careless errors. But six other subjects had ratios less than 10.

Insert Table 3 about here

The subjects in Study 2 were also asked, "How many pounds in a ton?" Their answers are shown in Table 4. Only 63% of the subjects gave the answer 2000 (one subject responded "5280").

Insert Table 4 about here

Study 3

Study 3 was designed to explore several issues concerning the results of the first two studies. First, we thought it was

Table 3

Ratio of Population of U.S. to Population of Oregon

Log_{10} of Ratio	Frequency of Responses Study 2	Study 3
> 6.0	1	1
5.0 to 5.9	6	4
4.0 to 4.9	11	6
3.0 to 3.9	28	23
2.0 to 2.9	48	42
1.0 to 1.9	59	38
No answer	9	5
Number of Subjects	170	128
% in "ballpark"[a]	37%	39%
Maximum	2,000,000	1,500,000
Third Quartile	1289	1200
Median	130	120
First Quartile	63	67
Minimum	.108	1

[a] 50 to 200, inclusive

Table 4

How Many Pounds in a Ton?

Response Range	Frequency
12-500	9
1000	12
1064-1600	6
2000	107 (63%)
2200-2800	11
3000-7000	10
10,000	4
100,000	6
1,000,000	2
No answer	3
Number of subjects	170

possible that subjects did not correctly understand the correspondence between numbers expressed as digits ("1,000,000") and expressed as words ("one million"). This might explain the subjects who estimated, for example, the population of the U.S. as "3,000,000,000."

Second, we introduced a simple aid designed to draw attention to the ratio between the two population estimates, hoping that the subjects could thereby improve their estimates.

Third, we explored the accuracy of estimation for several other items of general knowledge.

There were 128 subjects who were recruited, run, and paid as in Study 2.

Each subject received one of six questionnaires. One such is shown in Table 5. The others differed only in the first task, where the words "ten thousand" were replaced by "one hundred thousand," "ten million," "one hundred million," "ten billion," or "one hundred billion."

Insert Table 5 about here

The subjects' performance on the translation of number words into digits was excellent. Of 128 subjects, 7 failed to respond, 6 gave wrong answers, and 115 (90%) were correct. This performance rules out the possibility that this kind of confusion accounts for any appreciable proportion of population misestimates.

Table 5

Questionnaire for Study 3

Name________________________________ Sex M F Age____ (NOT DATE!)

Here is a number expressed in words. Under it please write the same number expressed with digits, with the appropriate number of zeroes and commas:

ten thousand

____

How many feet in a mile? __________________________

How many ounces in a pound? _______________________

What is the air distance, in miles, from Eugene to New York City? ____________

How many miles is it, by freeway, from Eugene to Portland? __________________

How many cups in a quart? __________________________

What is the population of the United States?

____

What is the population of Oregon?

____

Now compare your last two answers (population of US and Oregon). Test them against each other for consistency. To do so, form a ratio of the two answers, like this: $\frac{\#US}{\#OR}$

Form your ratio here: ______________________________

Simplify the ratio by lopping off zeroes from the top and bottom.

For example, $\frac{100}{10} = \frac{10\cancel{0}}{1\cancel{0}} = \frac{10}{1} = 10$. Further simplify the ratio as you can.

Now ask yourself, is this a sensible ratio? For example, suppose you end up with the ratio: $\frac{3}{2}$. This means that your estimates imply that for every two people in Oregon, there are only <u>three</u> people in the whole US, including Oregon. This is not sensible. After examining your ratio for sensibleness, please make two new estimates so that both of the estimates are sensible and their ratio is also sensible. If you are still satisfied with both your original estimates, just write "same" below.

What is the population of the United States?

____

What is the population of Oregon?

____

The answers to the measurement questions are shown in Table 6. Taken in conjunction with the pounds-in-a-ton results of Table 4, there is a mild suggestion in the data that large numbers (2000, 5280) are less well known or remembered than small numbers (4, 16).

Insert Table 6 about here

Results of the distance questions are shown in Table 7. The correct answer to the Eugene-Portland question is uncertain; depending on how the measurement is taken, the answer is from 108 to 120 miles. A "ballpark" estimate between 100 and 140 miles was given by 75% of the subjects.

Insert Table 7 about here

The air distance from Eugene to New York City is also not precisely known to us. An almanac lists the air distance from Portland to New York as 2454; the distance from Eugene (which is some 100 miles south of Portland) must be about the same. The slightly higher median of the subjects' reponses may reflect a confusion of air distance with highway distance (which the almanac shows, for Portland to New York, as 2992 miles). A "ballpark" estimate between 2000 and 3000 miles, inclusive, was given by only 51% of the subjects.

Tables 1 and 2 show the results from the first U.S. and

Table 6

Responses to Measurement Questions

Feet in a Mile		Ounces in a Pound		Cups in a Quart	
Range	Freq.	Range	Freq.	Range	Freq.
12-80	4	4	4	2	4
144-984	10	10	1	3	1
1000-1800	10	12	4	4	100 (78%)
2400-4950	14	16	114 (89%)	5	1
5000	6	32	3	6	2
5126-5275	10	45	1	8	13
5280	42 (33%)	64	1	12	1
5284-5860	14	168	1	16	1
6032-9600	4	Blank	2	32	1
10,000-25,000	3			42	1
100,000	1			Blank	3
Blank	10				
No. of Subjects	128		128		128

Table 7

Responses to the Distance Questions

Eugene-Portland by Freeway Range	Freq.	Eugene-NYC Air Distance Range	Freq.
2-90	3	300-700	5
100-107	28	1500-1802	7
108-115	38	2000	6
120	20	2100-2400	9
121-140	10	2500	7
150-196	17	2521-2950	11
200-300	8	3000	32
500	1	3025-3700	24
Blank	3	4000-8000	17
		10,000-15,000	3
		77,000	1
		Blank	6
% 100 - 140, incl.	75%	% 2000-3000, incl.	51%
First Quartile	108		2500
Median	114		3000
Third Quartile	130		3282

Oregon population questions and Table 3 shows the ratios of these two reponses. In these results the subjects of this study were highly similar to the subjects in Study 2. Although none of these subjects estimated the U.S. population as smaller than the Oregon population, one subject estimated them as the same, expressing the first as "250 million" and the second as "250,000,000." This 20-year-old female correctly expressed, at the top of the questionnaire, the words one hundred thousand as "100,000."

The instructions then asked the subjects to form and simplify the ratio of the U.S. population to the Oregon population. Of the 128 subjects, 6 did not complete this part, 9 formed the ratio but did not simplify it, 27 made errors in the simplification (e.g., 2 billion to 80 thousand was simplified to 2/8), and 86 (67%) performed this task adequately. The scoring criteria for this adequacy were lenient. Approximations (e.g., 420 million/4 million = 100), awkward fractions (e.g., 1 billion/1.5 million = 200/.3), and fractions not fully simplified (e.g., 250/2) were all coded as adequate.

The subjects then indicated new estimates, if they wished. Only 32 subjects (25%) changed one or both estimates, for a total of 41 changes. Of these, 21 changes made the estimate more accurate. Sometimes the change made both the estimate(s) and the ratio worse:

250 million / 3 million = 83 became

300 million / 5 million = 60.

Sometimes the subject changed the wrong estimate, so that the estimate became worse but the ratio improved:

2.4 billion / 68 thousand = 3529 became

2.4 billion / 68 million = 35.

Only 11 subjects made changes that improved the estimates(s) and the ratio. Some of them were just what we were looking for:

250 million / 40 million = 6 became

250 million / 4 million = 63.

Others, while improved, were still off the mark:

3 billion / 200 thousand = 15,000 became

3 billion / 300 thousand = 10,000.

Subjects who adequately formed and simplified the ratio were no more likely to make changes than those whose ratios were wrong or unsimplified.

Of the 40 subjects who gave U.S. estimates of one billion or more, only 6 (15%) made changes in these estimates. A listing of these changes indicates the flavor of the data:

1 billion	became	2 billion
2.5 billion	became	400 million
3 billion	became	2 billion
3 billion	became	10 million
5 billion	became	5 million
25 billion	became	100 billion

In short, this manipulation, asking subjects to re-evaluate their population estimates after considering whether the ratio of their two estimates was "sensible," produced only few changes,

with the changes as often for the worse as for the better.

Discussion

This series of three studies indicates that the subjects who volunteer for our experiments frequently lack knowledge of what can be considered basic facts, such as the population of the U.S. or the number of feet in a mile. A mild manipulation intended to persuade the subjects to re-evaluate and improve their estimates had little effect.

We see no reason to suppose that these subjects are in any significant way different from typical college students. And we suppose that college students are not below the national average in knowledge (although it could be argued that in another 20 or 30 years these subjects will accumulate a greater fund of mundane facts). Thus we are willing to generalize, quite tentatively, these results to broader populations.

When so generalized, these data suggest that efforts to develop estimation aids based on algorithmic decomposition techniques may fail not because people are unable to make such decompositions but because they have and hold extensive stores of misinformation. When this misinformation is brought to bear on estimation problems, degredation of performance, rather than improvement of performance, might be expected.